

PAPER DETAILS

TITLE: Kizilgerdan Kus Popölasyonu Biyoakustik Kayitlarinin Takibi İçin Kus Sesi Tanima Yöntemi
Gelistirilmesi

AUTHORS: Selim ARAS,Seda ÜSTÜN ERCAN

PAGES: 1720-1731

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/1177916>



Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Kızılgerdan Kuş Popülasyonu Biyoakustik Kayıtlarının Takibi İçin Kuş Sesi Tanıma Yöntemi Geliştirilmesi

 Selim ARAS ^{a,*},  Seda ÜSTÜN ERCAN ^a

^a Elektrik Elektronik Mühendisliği Bölümü, Mühendislik Fakültesi, Ondokuz Mayıs Üniversitesi, Samsun, TÜRKİYE

* Sorumlu yazarın e-posta adresi: selim.aras@omu.edu.tr
DOI : 10.29130/dubited.569642

ÖZET

Bu çalışmada, Kızılgerdan kuş popülasyonuna ait dört alt türün biyoakustik özelliklerinden tespiti için uygun öznelik ve sınıflandırma yöntemi araştırılmıştır. Özneliklerin belirlenmesi için Mel Frekansı Kepstrum Katsayıları temel alınmış ve bu katsayılardan istatistiksel parametreler yardımıyla hesaplanabilecek uygun öznelik araştırması yapılmıştır. Sınıflandırma aşamasında Doğrusal Ayırma Ayırıcı, Destek Vektör Makineleri ve k-En Yakın Komşuluk ve Ardışıl İleri Yönlü Öznelik yöntemleri kullanılmıştır. Sınıflandırıcı parametreleri 10-kat çapraz doğrulama yöntemi ile eğitim setinde belirlenmiştir. Daha sonra, eğitilmiş sınıflandırıcı parametreleri test veri setine uygulanarak sınıflandırma doğruluğu elde edilmiştir. Sonuç olarak, çalışmamızda Mel Frekansı Kepstrum katsayıları temel alınarak hesaplanan ortalama, etkinlik ve karmaşıklık parametreleri k-En Yakın Komşuluk Yöntemi ile sınıflandırıldığında en iyi başarımlar elde edilmiştir. Önerdiğimiz yöntemin sınıflandırma başarımları eğitim kümesinde %97, test kümesinde ise %94 olarak elde edilmiştir.

Anahtar Kelimeler: Kuş sesi sınıflandırma, Mel frekansı kepstrum katsayısı, İstatistiksel parametreler

Developing Bird Song Recognition Method for Monitoring Robin Birds Population Bioacoustics Records

ABSTRACT

In this study, suitable features and classification methods were investigated to determine the four subspecies of Robin birds population from their bioacoustic characteristics. Mel Frequency Cepstrum Coefficients were taken as basis for the determination of the features and a suitable feature search was performed by using statistical parameters from these coefficients. In the classification stage Linear Discriminant Analysis, Support Vector Machines, k-Nearest Neighborhood, and Sequential Forward Feature Selection methods were used. Classifier parameters were determined by 10-fold cross validation method. Then, the classification accuracy was obtained by applying the trained classifier parameters to the test data set. As a result, in our study, the best performance was obtained when the mean, efficiency and complexity parameters, which were calculated based on Mel Frequency Cepstrum coefficients, were classified by k-Nearest Neighborhood Method. The classification performance of the proposed method was obtained 97% in the training set and 94% in the test set.

Keywords: Bird song classification, Mel frequency cepstrum coefficients, Statistical parameters

I. GİRİŞ

Kuşlar, ekolojik dengeyi sağlamada ve besin çemberinde önemli rol oynamaktadır. Ancak, insan kaynaklı iklim değişikliği ve habitat tahribatı nedeniyle birçok kuş türünün nüfusu azalmaktadır. Bu olumsuz tablo, çeşitli koruma çabalarına yol açmıştır. Bunlar arasında kuş çeşitliliğinin belirlenmesi ve belirli türlerin popülasyon kayıtlarının tutulması yer almaktadır [1]. Kuşların doğal ortamlarında elle incelenmesi zaman alıcı ve pahalıdır. Ayrıca deneyimli kuş gözlemcisi gereklidir. Bu nedenle, kuşları doğal ortamlarında araştırmak için otomatik yöntemlerin geliştirilmesine ihtiyaç duyulmaktadır. Kuş seslerinin otomatik tanınması çalışmaları, yaban hayatının izlenmesi için önemlidir. Çünkü birçok biyolojik izleme projesi kuşların akustik olarak tespit edilmesine dayanmaktadır. Bu amaçla oluşturulmuş, dünyanın farklı köşelerinden kuş ses kayıtlarının yüklendiği ve etiketlendiği önemli erişilebilir biyoakustik bilgi kaynakları mevcuttur. Bunlardan en önemlilerinden biri de bu çalışmada kullanılan Xeno-Canto [2] veri setidir.

Literatürde Xeno-Canto veri setindeki kuş seslerini kullanan birçok çalışma bulunmaktadır. Bu çalışmalardan; Kızılgerdan kuşu seslerinin zamansal, spektral ve yapısal özellikleri incelendiği bir çalışmada [3] Doğrusal Öngörülü Kodlama (DÖK) temelli çerçeve seviyesi farklılıklarının ortalaması olarak tanımlanan bir mesafe ölçüsüne göre karşılaştırma yapılarak kayıtlardan kuş sesleri kümelenerek ayrılmıştır. Kümelenme sonuçlarından elde edilen sesler, Gizli Markov Modeli (GMM) tabanlı Kızılgerdan sesi algılayıcısının akustik modellemesini geliştirmek için kullanılmıştır. Başka bir çalışmada [4] ise Hızlı Fourier Dönüşümü (HFD) ve DÖK'e dayalı öznelilikleri kullanılarak dinamik zaman bükme algoritması ile GMM'nin kuş seslerinin otomatik tanınmasındaki başarısı araştırılmıştır. Bunun dışında kuşların görüntü ve seslerine göre sınıflandırma yapan bir yöntem de önerilmiştir [5]. Kuş sesleri için Xeno-Canto veri seti kullanılmış, görüntülerde ölçekli değişmeyen özellik dönüşümü, kuş seslerinde ise Mel Frekansı Kepstrum Katsayılarını (MFKK) öznelik olarak kullanarak destek vektör makinesi ile bu öznelikler sınıflandırılmıştır. McIlraith ve Card [6], kuş seslerinin sınıflandırılması için DÖK ve HFD ile birlikte geri yayımlı iki katmanlı ileri beslemeli bir sinir ağı kullanılması yöntemini önermişlerdir. Kuş seslerinin sınıflandırılması için temel bileşenleri ve ikinci dereceden diskriminant analizini içeren çok değişkenli istatistiklerden gelen yöntemleri ve geri yayımlı iki katmanlı ileri beslemeli bir sinir ağı kullanılması da önerilmiştir [7]. Başka bir çalışmada [8], MFKK katsayıları öznelik olarak seçilmiş ve üç katmanlı derin sinir ağının yanında Gaus Karışım Modeli ve Destek Vektör Makinesi (DVM) yöntemleri kullanılmıştır. Alborno ve ark. [9] ise yerel bir bölgedeki kuş türünü sesinden tanımak için MFKK ile birlikte zaman ve frekans analizinde kullanılan parametreleri öznelik olarak kullanmıştır. Sınıflandırma aşamasında ise çok katmanlı algılayıcılar, DVM ve Rassal Orman Algoritması (Random Forest Algorithm) tekniklerinden yararlanmıştır. Priyadarshani ve ark., dalgacık düğümlerinin bir alt kümesinden sesin yeniden yapılandırılmasına dayanan, yeni türler üzerinde hızlı ve kolay bir şekilde eğitilebilecek, sürekli saha kayıtları için bir kuş çağrı sesi algılama yöntemi önermiştir. Çalışmalarında Rocky Dağı Biyoloji Laboratuvarına (Rocky Mountain Biological Laboratory -RMBL) ait Amerika Kızılgerdanları veri setini de kullanmışlardır [10].

Akustik iletişim kuşlarda çok zengin ve sürekli [11]. Bu süreklilik sayesinde, ses analiz yöntemlerini kullanarak kuş popülasyonlarını tespit etmek mümkündür. Günümüzde kuş çeşitliliğini izlemek için kullanılabilecek çok miktarda akustik veri kaydedilmektedir. Programlanabilir kayıt cihazları ve ses tanıma teknolojisindeki ilerlemeler sayesinde kuş türlerinin otomatik olarak tanımlanması mümkündür. Bu çalışmada, Xeno-Canto açık veri tabanında yer alan Kızılgerdan kuşlarının dört alt türü için ses tanıma çalışması yapılmıştır. Siyah Kenarlı Kızılgerdan, Gri Başlı Kızılgerdan, Beyaz Kaşlı Çalı Kızılgerdanı ve Japon Kızılgerdanı kuşlarının her biri için 50'şer olmak üzere toplamda 200 adet ses kaydı kullanılmıştır. Ses kayıtlarından öznelik elde etmek için MFKK'dan hesaplanan istatistiksel parametrelerden faydalanılmıştır. Test aşamasında Doğrusal Ayırma Analizi (DAA), Destek Vektör Makinesi (DVM) ve k-En Yakın Komşuluk (k-EYK) kullanılmıştır. Elde edilen sonuçlar Ardışıl İleri Yönlü Öznelik Seçme (AİYÖS) yöntemi ile geliştirilmiştir.

II. MATERYAL YÖNTEM

A. VERİ SEÇİMİ

Xeno-canto kuş seslerini paylaşmaya adanmış, 10.000'e yakın kuş türü için yaklaşık 400.000 kayıt içeren açık veri tabanı sunan bir web sitesidir [12]. Kullanıcılar kendi kayıtlarını yükler ve bunları cins, tür, alt tür, yer, kalite (A'dan E'ye, A sesin en iyi kalitesidir) vb. ile etiketler. Sesler ayrıca “çağrılar” veya “şarkılar” olarak da sınıflandırılabilir.

Bu çalışmada, Siyah Kenarlı Kızılgerdan (*Poecilodryas hypoleuca*), Gri Başlı Kızılgerdan (*Heteromyias cinereifrons*), Beyaz Kaşlı Çalı Kızılgerdanı (*Tarsiger indicus*) ve Japon Kızılgerdanı (*Larvivora akahige*) kuşlarına ait sesler kullanılmıştır. Kuş sesleri, kayıtlar içerisinde başlangıç ve bitiş noktaları seçilerek elde edilmiştir. Her alt tür için 50'şer olmak üzere toplamda 200 adet ses kaydı kullanılmıştır. Arka plan gürültüsü yüksek olduğundan kuş şarkıları yalnızca A derecesini içerecek şekilde seçilmiştir.

B. ÖZNİTELİK ÇIKARMA YÖNTEMİ

Çalışmamızda Kızılgerdan kuş seslerinin sınıflandırılmasında MFKK katsayılarından elde edilen istatistiksel parametrelerin öznitelik olarak kullanılması yöntemi önerilmiş ve başarımları irdelenmiştir. Kuş seslerinin başlangıç ile bitiş noktaları seçilerek belirlendiğinden ve süreleri kuşlara göre değişkenlik gösterdiğinden veri setinde kullandığımız kuş ses kayıtları farklı uzunluklara sahiptir. MKKK ise sesin akustik özelliklerinin durağan kabul edildiği kısa çerçevelere uygulanır [13],[14]. Çerçeve uzunluğu sabit tutulduğundan kayıtlardan elde edilen toplam çerçeve sayılarında farklılıklar oluşmaktadır. Her ne kadar çerçevelerden 12'şer adet MFKK katsayısı hesaplanmış olsa da sonuçta bir kuş sesinden elde edilen öznitelik matrisinin boyutları değişken çerçeve sayısı nedeniyle farklı olmaktadır. Bu durum ise sınıflandırma aşamasında eşit sayıda öznitelik kullanan sınıflandırma yöntemlerini devre dışı bırakmaktadır. Hem değişken süreli kayıtlardan hesaplanan MFKK katsayılarından eşit sayıda öznitelik elde etmek, hem de özniteliklerin sınıfları ayırt etmedeki yeteneğini artırmak için bu çalışmada istatistiksel parametrelerden faydalanılmıştır. Bir kuş ses kaydından MFKK kullanılarak hesaplanan öznitelik matrisinden istatistiksel parametrelerin nasıl hesaplandığı Şekil 1'de gösterilmektedir.

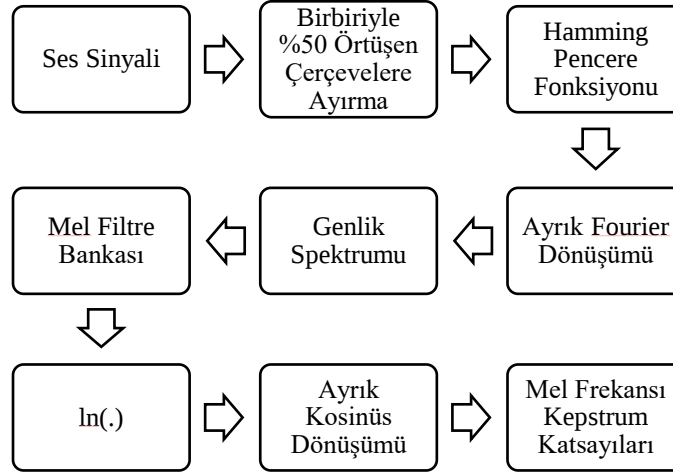
1.Çerçeve için MFKK	m_1^1	m_1^2		m_1^{11}	m_1^{12}
2.Çerçeve için MFKK	m_2^1	m_2^2		m_2^{11}	m_2^{12}
⋮	⋮	⋮	⋮	⋮	⋮
(N-1). Çerçeve için MFKK	$m_{(N-1)}^1$	$m_{(N-1)}^2$		$m_{(N-1)}^{11}$	$m_{(N-1)}^{12}$
N. Çerçeve için MFKK	m_N^1	m_N^2		m_N^{11}	m_N^{12}
İstatistiksel Parametre	p_1	p_2		p_{11}	p_{12}

Şekil 1. MFKK katsayılarından parametrelerin hesaplanması süreci

C. MFKK

Ses sinyali öncelikle birbirleriyle %50 örtüşen N adet çerçevelere ayrılır. Daha sonra her bir çerçeveden 12 adet MFKK hesaplanır. Birinci çerçeve için hesaplanan MFKK katsayıları öznitelik matrisinin birinci satırını oluşturur. Daha sonra, her çerçeve için hesaplanan MFKK katsayıları

öznitelik matrisinin yeni bir satırını oluşturur. Sonuç olarak Nx12 boyutunda bir öznitelik matrisi elde edilir. Böylece her çerçeve için MFKK katsayılarının birincisinden sonuncusuna kadar hesaplanan katsayılar aynı sırada alt alta dizilmiş olur. Bundan sonraki adım, öznitelik matrisinin her sütunu için ayrı ayrı istatistiksel parametrenin hesaplanmasıdır. Ses kaydının uzunluğuna bağlı olarak çerçeve sayısı (N) değişkenlik gösterse de bu istatistiksel parametrelerin hesaplanması sonucunda her ses kaydı için öznitelik olarak 12 elemanlı bir vektör elde edilir. Bu durum sınıflandırma aşamasında sınıflandırma yöntemi seçiminde esnekliği sağlamaktadır.



Şekil 2. MFKK hesaplama süreci akış diyagramı

MFKK elde edilme süreci için oluşturulan bir akış diyagramı şekil 2’de gösterilmektedir. Öncelikle ses sinyali bu çalışmada uzunluğu 20ms olarak belirlenen ve birbirleriyle %50 örtüşen çerçevelere ayrılmıştır. Daha sonra ise spektral sızıntıları önlemek için her çerçeveye denklem 1’de verilen Hamming pencere fonksiyonu uygulanır. Bu fonksiyonun sadece N_w kadar uzunluktaki bir aralıkta sıfır olmadığı var sayılır. Burada n , çerçevedeki örnek sayısıdır.

$$w[n] = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N_w}\right), & 0 \leq n \leq N_w - 1 \\ 0, & N_w - 1 < n < 0 \end{cases} \quad (1)$$

Enerji tayf yoğunluğu hesaplandıktan sonra mel filtre bankası oluşturulur. Bunun için frekans tayfındaki sınır frekans değerleri (f) denklem 2 yardımıyla mel frekansına dönüştürülür.

$$f_{mel} = 2595 \log_{10}\left(\frac{f}{700} + 1\right) \quad (2)$$

En küçük ve en büyük mel frekansı aralığındaki tayf, filtre bankasındaki filtre sayısına bağlı olarak eşit aralıklarla bölünür. Böylelikle filtrelerin sınırları belirlenmiş olur. Bu sınırlar ters dönüşüm ile mel frekansından tekrar normal frekansa dönüştürüldüğünde ise bu iki frekans birbirine göre doğrusal olmadığından dolayı mel filtre bankasındaki üçgen filtrelerin genişlikleri de farklı olur.

Bundan sonraki amaç ilgilenilen frekans bölgelerindeki genlik spektrumunu elde etmektir. Genlik spektrumu, i . çerçeve için denklem 3’ten hesaplanır.

$$G_i(k) = |X_i(k)| \quad (3)$$

Genlik spektrumu Mel filtre bankasından geçirildikten sonra logaritması alınır. Son aşamada MFKK katsayılarını elde etmek için denklem 4’te verilen ayırık kosinüs dönüşümü uygulanır. Sonuç ise Mel Frekansı Kepstrum Katsayıları olarak adlandırılır.

$$c(l) = \sum_{m=1}^M X_i(m) \cos \left(l \frac{\pi}{M} \left(M - \frac{1}{2} \right) \right) \quad (4)$$

Burada M , $m = 1, 2, \dots, M$ ($M \ll N$) için filtre bankasındaki filtre sayısıdır.

D. İSTATİSTİKSEL PARAMETRELER

Bu çalışmada istatistiksel olarak Hjorth tanımlayıcıları (etkinlik, hareketlilik ve karmaşıklık) ile aritmetik ortalama, standart sapma, basıklık ve çarpıklık parametreleri kullanılmıştır.

Etkinlik Parametresi: Zaman fonksiyonun varyansı (standart sapmanın (σ) karesi) olarak ifade edilen etkinlik parametresi, sinyalin ortalama enerjisidir ve denklem 3'deki gibi gösterilir.

$$H_E = \sigma_0^2 \quad (5)$$

Hareketlilik: Hareketlilik, sinyalin güç spektrumunun standart sapmasının oranını ifade eder. Sinyalin birinci türevinin etkinlik parametresinin, sinyalin etkinlik parametresine oranının karekökü olarak denklem 4'de de gösterildiği gibi hesaplanmaktadır.

$$H_H = \frac{\sigma_1}{\sigma_0} \quad (6)$$

Karmaşıklık: Frekanstaki değişimi gösteren karmaşıklık, sinyalin birinci türevinin hareketliliği ile sinyalin hareketliliği arasındaki oran olarak denklem 5'deki gibi hesaplanır.

$$H_K = \sqrt{\left(\frac{\sigma_2}{\sigma_1} \right)^2 - \left(\frac{\sigma_1}{\sigma_0} \right)^2} \quad (7)$$

Aritmetik Ortalama: Bir veri setindeki verilerin değerlerinin toplamının veri sayısına bölünmesi ile elde edilir ve denklem 6'dan faydalanılarak hesaplanır.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (8)$$

Standart Sapma: Veri setindeki bulunan her bir değerin birbirine olan yakınlığını ve uyumunu gösteren istatistiksel bir değişkendir ve denklem 7'deki gibi hesaplanır.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (9)$$

Basıklık: Veri değerlerinin, eğri grafiği üzerinde meydana getirdiği düzlüğün veya sivrilğin bir göstergesidir ve denklem 8'deki gibi hesaplanır.

$$K = \frac{1}{N} \sum_{i=1}^N \left[\frac{x_i - \bar{x}}{\sigma} \right]^4 \quad (10)$$

Çarpıklık: Veri setindeki örneklerin ortalama çevresindeki dağılımına bakmaktadır. Veriler ortalamaya göre sağa doğru veya sola doğru çarpık olabilirler. Eğer çarpıklık yoksa eğri normal dağılıma sahiptir denilebilir. Çarpıklık katsayısı denklem 9'daki gibi hesaplanır.

$$S = \frac{1}{N} \sum_{i=1}^N \left[\frac{x_i - \bar{x}}{\sigma} \right]^3 \quad (11)$$

E. SINIFLANDIRMA

Sınıflandırma bir veri setindeki verilerin özelliklerine bakılarak, veriyi gruplara ayırma işlemidir. Öznitelikler AİYÖS yöntemine göre gruplandırıldıktan sonra önerilen yöntemin sınıflandırmadaki başarımını test etmek için 10-kat çapraz doğrulama uygulanmıştır. 10-kat çapraz doğrulamada eğitim setine ait denemeler 10 eşit alt kümeye bölündükten sonra her defasında 1 alt küme onaylama kümesi olarak dışarıda bırakılır. Kalan 9 küme en yüksek sınıflandırma başarımının yakalandığı durumdaki sınıflandırma parametresini hesaplamak için eğitim kümesi olarak kullanılmıştır. Sınıflandırma parametresinin en iyi değerine karar verebilmek için bu işlem 50 kez tekrarlanmıştır. Eğitim kümesinden elde edilen en iyi sınıflandırma parametresi değeri ise daha sonra test kümesini sınıflandırmak için kullanılmıştır. Bu çalışmada verileri sınıflandırmak için aşağıda belirtilen 3 yöntemden faydalanılmıştır.

k-En Yakın Komşuluk Yöntemi: Bu yöntemde bir verinin sınıflandırılması, öğrenme verisinde sınıfı daha önceden bilinen en yakın k -adet komşunun sınıfına bağlıdır. Belirlenen komşuların en fazla temsil ettiği sınıf test verisinin sınıfı olarak atanır [15]. $y_i \in \{1, \dots, c\}$ sınıf etiketlerini, $L = \{(y_i, x_i), i = 1, \dots, n_L\}$ ise eğitim setini temsil etsin. Öznitelik vektörleri $x_i \in \mathbb{R}^p$ olmak üzere, sınıfı belirlenecek bir x test verisi için ilk önce eğitim kümesinin en yakın elemanı (y_{i1}, x_{i1}) belirlenir. $i_1 = \operatorname{argmin}_i (d(x, x_i))$ için denklem 12'deki gibi Öklid uzaklığı tanımlanır.

$$d(.,.)d(x, x_i) = \sqrt{(x - x_i)^T (x - x_i)} \quad (12)$$

Bu durumda test verisi sınıflandırma sonuç etiketi ise y_{i1} 'dir. Genellikle, L 'den sadece bir gözlem yerine, k -adet en benzer elemanlar göz önünde bulundurulur. Bu nedenle, $Y = \{y_{i1}, \dots, y_{ik}\}$ için sınıf etiketi sayıları her bir $K_i = \sum_{y \in Y} \delta_{iy}$ sınıfı için belirlenir, burada δ_{iy} , dürtü işlevi anlamına gelir. Sınıf etiketi, $Y = \operatorname{argmax}_i \{K_1, \dots, K_c\}$ sonuçlarında bulunan en yaygın sınıf olarak belirlenir.

Doğrusal Ayırıştırma Analizi: Bu yöntemde farklı sınıf verilerinin iz düşümlerini ayıracak bir doğru aranır. Ayrıntıları [16]'de ayrıntılı olarak sunulan DAA, verileri p' boyutlu $\mathbb{R}^{p'}$ uzayından, daha düşük boyutlu $\mathbb{R}^d$ sınıflandırma uzayına dönüştüren optimal bir doğrusal sınıflandırma fonksiyonu üreten denetimli bir tekniktir. Sınıflar arası dağılım matrisi denklem 13'deki gibi hesaplanır.

$$B = \frac{1}{k-1} \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (13)$$

Burada $\bar{x}$, $\bar{x} = \frac{1}{k} \sum_{i=1}^k \bar{x}_i$ sınıfının ortalamasını belirtir. n_i ise i . sınıfındaki örneklerin sayısıdır. Sınıf içi dağılım matrisi denklem 14'deki gibidir.

$$W = \frac{1}{n-k} \sum_{j=1}^k \sum_{x_i \in C_j} (x_i - \bar{x}_j)(x_i - \bar{x}_j)^T \quad (14)$$

Burada n , tüm sınıflardaki örneklerin toplam sayısıdır. Kendi özdeğerlerine göre sıralanan $W^{-1}B$ matrisinin özvektörlerine kanonik vektörler denir. Birinci d kanonik vektörlerin seçilmesi ve bunların $\tilde{A}^{(d)} \in \mathbb{R}^{d \times p'}$ projeksiyon matrisi olarak satır sıralanmasıyla herhangi bir $x \in \mathbb{R}^{p'}$ vektörü, bir alt-boyutlu öznitelik alanı $\mathbb{R}^d$ üzerine yansıtılabilir. k -sınıflı bir problem için LDA basit ve etkili bir şekilde uygulanabilir. Eğer $\|\tilde{A}^{(d)}(x_i - \bar{x}_j)\| < \|\tilde{A}^{(d)}(x_i - \bar{x}_k)\|$ eşitsizliği gözlemlenirse x vektörü C_j sınıfına sınıflandırılır. Tüm $k \neq j$. $\|\cdot\|$ için Öklid normunu belirtir.

Destek Vektör Makineleri: Destek Vektör Makineleri, temel olarak iki sınıfa ait verileri birbirinden ayırmak için hiper düzlemleri kullanır. Destek Vektör makineleri (SVM) aşağıdaki gibi tanımlanabilir [17]. $L = \{(y_i, x_i), i = 1, \dots, n_L\}$, $x_i \in \mathbb{R}^p$ doğrusal olarak ayrılabilen eğitim örnekleri kümesidir. Burada $y_i \in \{-1, 1\}$ sınıf etiketlerini gösterir. $\mathbb{R}^p$ 'de x 'i ayıran, $w \cdot x + b = 0$ olan p boyutlu bir hiper düzlem varsayılır. Ayrılan $|w \cdot x + b| = 1$ ve $|w \cdot x + b| = -1$ hiper düzlemler arasındaki mesafe

$\frac{2}{\|w\|}$, dir. Optimum ayırıştırma hiper düzlemi, tüm $x_i, i = 1, \dots, n_L$ değerleri için, $y_i(x_i \cdot w + b) \geq 1$ sınırları dahilinde $\frac{\|w\|^2}{2}$, nin minimize edilmesiyle bulunabilir.

Veriler doğrusal olarak ayrılmadığında, sınıflandırma hatalarıyla orantılı bir miktarı en aza indirirken, kenar boşluğunu maksimize eden bir hiper düzlem, bir sınırlama ile sonuçlanan pozitif durgunluk değişkenleri ξ_i eklenerek belirlenir: $y_i(x_i \cdot w + b) \geq 1 + \xi_i$ ve $\min(w) = \frac{\|w\|^2}{2} + C \sum_{i=1}^n \xi_i$. Burada C , kullanıcı tarafından seçilen, kenar genişliği ve yanlış sınıflandırma hataları arasındaki dengeyi kontrol eden bir ceza faktörüdür.

Karar fonksiyonu doğrusal olmadığında, daha yüksek boyutlu bir Öklid uzayına H verilerin ilk eşlemesi $\phi: \mathbb{R}^{n_L} \rightarrow H$ olarak gerçekleştirilir. Böylece sınıflandırma sorunu yeni uzayda formüle edilir. Bu durumda, eğitim algoritması sadece $\phi(x_i) \cdot \phi(x_j)$ 'in H formunda nokta çarpımındaki verilere bağlıdır. Mercer teoremi [18], $\phi(x_i) \cdot \phi(x_j)$ 'nin pozitif bir kesin simetrik çekirdek işlevi $K(x_i, x_j)$ ile değiştirilmesine izin verir. Örneğin $\gamma > 0$ için Gauss radyal-temel işlevi $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ 'dir.

F. ARDIŞIL İLERİ YÖNLÜ ÖZNİTELİK SEÇME YÖNTEMİ (AİYÖS)

Bir veriyi çok sayıda farklı öznitelikler ile temsil etmek mümkündür. Bu özniteliklerin sayısı arttıkça hem veriyi işlemek için gereken süre artmakta hem de bazı durumlarda uygulamaların başarımı düşmektir. [19] Bu yöntemin amacı en yüksek doğrulama başarımını gösteren öznitelikleri çapraz doğrulama yöntemi ile belirlemektir. Bunu da teker teker seçim kümesine ekleyerek belirlemektedir.

Basit olmasına rağmen çokça tercih edilen bir yöntemdir. AİYÖS, öncelikle özniteliklerin ayrı ayrı sınıflandırma başarımlarının en yüksekini arar. Daha sonra en yüksek başarıma sahip öznitelik ile diğer öznitelikler ikili olarak birleştirilir ve yine en yüksek başarımlar aranır. Başarım yüzdesi mevcut başarımdan daha düşük olmadıkça, özniteliklerin grup olarak birleştirilmelerine devam edilir. Bu süreçte AİYÖS ile elde edilen en yüksek başarımlar, sınıflandırma sonucu olarak belirlenir.

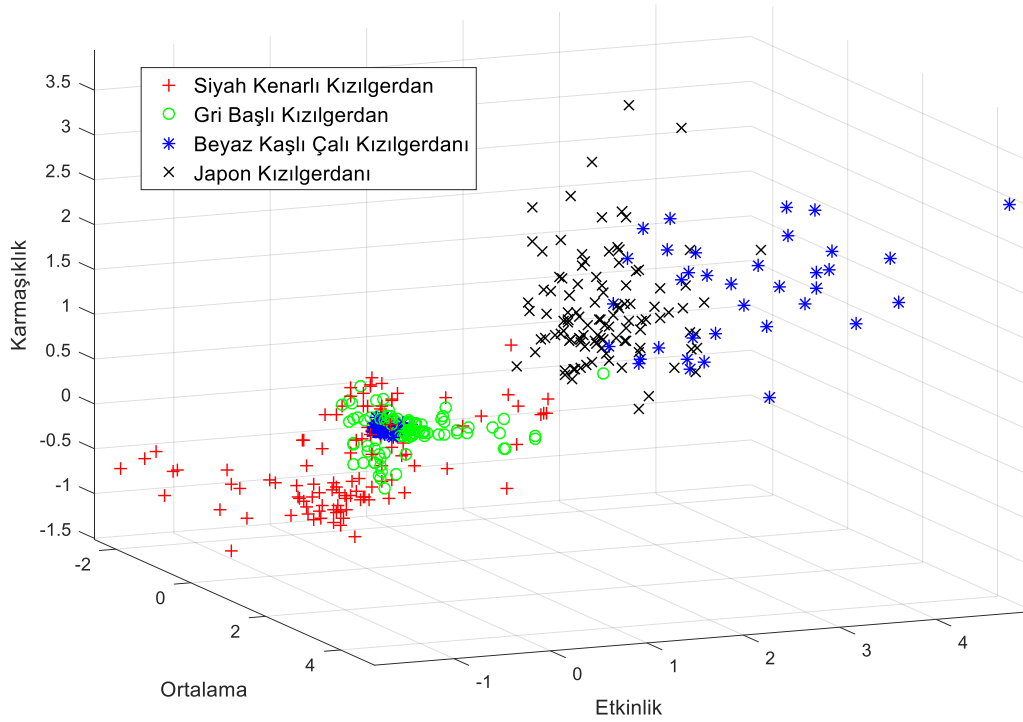
III. SONUÇ

Bu çalışmada, ses kayıtlarını kullanarak kuş türlerinin tanımlanması için MFKK ve istatistiksel parametre tabanlı bir yöntem önerilmiştir. MFKK'nın ham katsayılarının yanında MFKK'dan parametreler yardımıyla elde edilen özniteliklerin iyi bir alt tür ayrımcılığı sağladığı ve sınıflandırma çerçevesinde bir özellik gösterimi olarak kullanılabileceği gösterilmiştir. Bunun dışında, sınıflandırma başarımlarını artırmak için özniteliklerin ayırıştırma yeteneğini artırmak için AİYÖS yöntemi de kullanılmıştır.

Tablo 1. Eğitim ve test kümesindeki sınıflandırma sonuçları

<i>Eğitim Kümesindeki Sınıflandırma Sonuçları</i>												
	<i>AİYÖS 1.Adım</i>			<i>AİYÖS 2.Adım (Ortalama + diğer öznitelikler)</i>			<i>AİYÖS 3.Adım (Ortalama + Etkinlik + diğer öznitelikler)</i>			<i>AİYÖS 4.Adım (Ortalama + Etkinlik + Karmaşıklık + diğer öznitelikler)</i>		
	<i>DAA</i>	<i>DVM</i>	<i>k-EYK</i>	<i>DAA</i>	<i>DVM</i>	<i>k-EYK</i>	<i>DAA</i>	<i>DVM</i>	<i>k-EYK</i>	<i>DAA</i>	<i>DVM</i>	<i>k-EYK</i>
<i>Etkinlik</i>	73	74	76 (k=1)	86	96	96 (k=1)		sabit			sabit	
<i>Hareketlilik</i>	48	50	59 (k=3)	84	88	89 (k=1)	87	96	96 (k=1)	88	95	90 (k=1)
<i>Karmaşıklık</i>	56	59	58 (k=1)	82	93	88 (k=1)	89	95	97 (k=1)		sabit	
<i>Ortalama</i>	78	90	91 (k=3)		sabit			sabit			sabit	
<i>Standart Sapma</i>	75	76	78 (k=1)	89	95	93 (k=1)	87	96	94 (k=1)	88	94	91 (k=1)
<i>Basıklık</i>	46	43	48 (k=5)	81	85	87 (k=1)	88	95	91 (k=1)	88	96	92 (k=1)
<i>Çarpıklık</i>	53	57	50 (k=1)	88	90	86 (k=3)	91	95	90 (k=5)	90	96	88 (k=3)
<i>Test Kümesindeki Sınıflandırma Sonuçları</i>												
<i>AİYÖS ile belirlenen (Ortalama + etkinlik + karmaşıklık) öznitelikleri</i>							<i>DAA</i>	<i>DVM</i>	<i>k-EYK</i>			
							86	93	94 (k=1)			

Eğitim ve test kümesindeki sınıflandırma sonuçları Tablo 1’de verilmiştir. Her özneliğin ya da AİYÖS ile oluşturulan öznelik gruplarının sınıflandırma sonuçları üç farklı sınıflandırma algoritması ile test edilmiştir. Özneliklerin bireysel sınıflandırma sonuçları incelendiğinde; MFKK katsayılarının ortalama değerleri elde edilerek oluşturulan özneliğin %91 ortalama ile en iyi sonucu verdiği anlaşılmaktadır. Kendisinden sonra gelen standart sapma ile oluşturulan özneliğin ise sınıflandırma sonucu %13 daha azdır. AİYÖS yöntemine göre, ortalama özneliği sabit tutularak diğer öznelikler ile ikili gruplarının sınıflandırma sonuçlarına bakılmıştır. Burada ortalama ve etkinlik özneliği birlikte kullanıldığında, ortalamanın ilk sonucuna göre %5’lik bir iyileşme ile sınıflandırma sonucunun geliştirildiği görülmektedir. Her ne kadar standart sapma özneliğine ait sınıflandırma sonucunun ilk sonuçlara bakıldığında etkinlik özneliğine göre fazla olduğu görülse de AİYÖS yöntemi sınıfları en iyi şekilde ayırtıran öznelikleri arar. Bu durumda Ortalama ve standart sapma özneliklerinin daha çok aynı sınıfları ayırtırma yeteneği bulunurken en iyi ikili grubun ise farklı sınıfları ayırtırabilme yeteneğine sahip özneliklerden oluştuğu anlaşılmaktadır.



Şekil 3. MFKK hesaplama süreci akış diyagramı

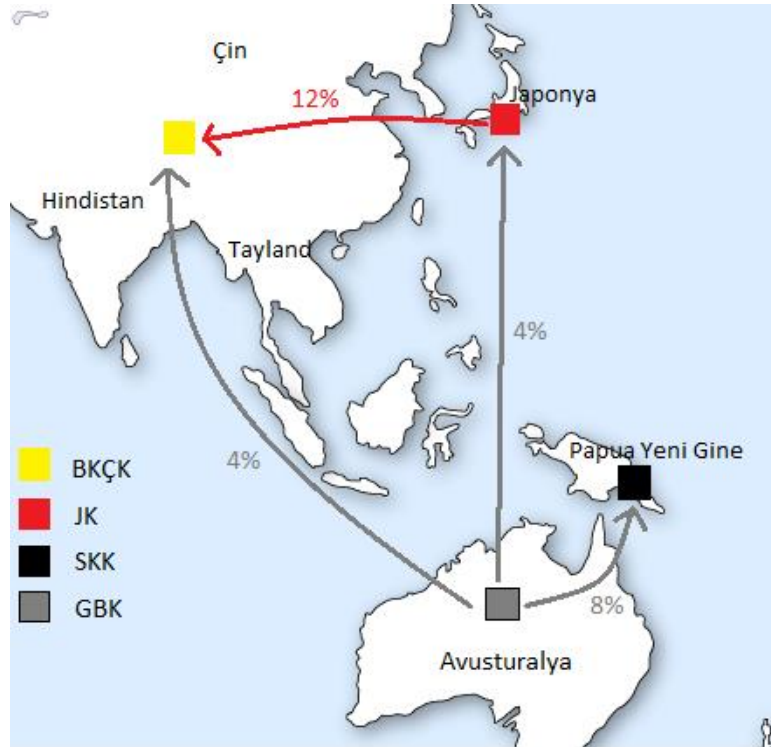
En iyi ikili grubu sabit tutarak kalan öznitelikler ile üçlü gruplarının sınıflandırma sonuçlarına baktığımızda ise ortalama ve etkinlik parametrelerine ek olarak karmaşıklık parametresi eklendiğinde eğitim kümesinden elde edilen sonucun %1 daha gelişerek % 97'ye ulaştığı görülmektedir. AİYÖS yöntemine göre bu üçlü öznitelik grupları sabit tutularak diğer öznitelikler ile dördümlü öznitelik gruplarının sonuçları irdelendiğinde ise en iyi sonucun en iyi üçlü öznitelik gruplarına göre %1 düşerek %96'ya gerilediği görülmektedir. Dolayısıyla AİYÖS yöntemine göre öznitelik gruplarının araştırma süreci son bulmuş olur. Sonuç olarak eğitim kümesi üzerinden en iyi sonucu veren öznitelik grubu k-NN (k=1) ile %97'lik bir başarımla ortalama, etkinlik ve karmaşıklık parametreleridir. En iyi öznitelik grubunu oluşturan bu üç parametre birbirlerine göre çizdirildiğinde sınıfların dağılımı ve özniteliklerin sınıfları ayırmadaki başarısı Şekil 3'de gösterilmektedir. Tablo-1'de de görüldüğü üzere, model olarak belirlenen bu üç öznitelik grubunun test kümesindeki sonucu %94 (k-NN (k=1)) olarak yüksek başarımla elde edilmiştir.

Gerçek sınıf	SK	100%				100%	
	GB	8%	84%	4%	4%	84%	16%
	BK			100%		100%	
	JK			12%	88%	88%	12%
		SKK	GBK	BKÇK	JK	Doğru pozitif oranı	Yanlış pozitif oranı
		Tahmin edilen sınıf					

Şekil 4. Test kümesinin hata matrisi

Sınıflandırma modellerinin performansını değerlendirmek için kullanılan yöntemlerden biri de hedef niteliğe ait tahminlerin ve gerçek değerlerin karşılaştırıldığı hata matrisinin incelenmesidir. En iyi sınıflandırma parametrelerine göre test kümesinin hata matrisi şekil 4’te gösterilmiştir.

Yeşil renk doğru pozitif oranı gösterirken, kırmızı renk ise yanlış pozitif oranını göstermektedir. Bu bilgiler ışığında, önerdiğimiz yöntemin Siyah Kenarlı Kızılgerdan ile Beyaz Kaşlı Çalı Kızılgerdan türlerini net olarak diğer türlerden ayırabildiği görülmektedir. Japon Kızılgerdanı ise %12’lik bir hata oranı ile Beyaz Kaşlı Çalı Kızılgerdanı türü ile karıştırılmıştır. Sesi diğer türlerden daha zor ayırt edilen Gri Başlı Kızılgerdan türü ise %16 ile en yüksek hata oranına sahiptir. Diğer üç türe de hatalı atamalar gerçekleşmiştir. Bu hatalı atamaların sebeplerinden biri de dört kuş türünün yaşadığı coğrafyanın yakınlık-uzaklık ilişkisidir. Bu coğrafi ilişki hata matrisindeki yüzdelere ile birlikte şekil 5’te sunulmaktadır.



Şekil 5. Kuş türlerinin yaşadığı coğrafi bölgeler ve test kümesindeki hatalı tahmin yüzdelere dağılımı

Bu çalışma, akustik özelliklere dayanarak kuş türleri tanınmasının mümkün olduğunu göstermektedir. Önerilen yaklaşım, akustik bilgilerin mevcut olduğu durumlarda MFKK ve istatistiksel parametrelerin öznel olarak kullanılmasının Kızılgerdan kuş türlerini ayırmada yüksek başarıma ulaşabildiğini göstermiştir. Gelecekteki çalışma, kuş türlerinin gerçek zamanlı sınıflandırılmasına yönelik olacaktır. Sadece Kızılgerdan kuşlarının alt türleri ile çalışılacak olup, sınıflandırmada Kızılgerdanı temsil etmeyen sesler için diğer sınıfı da oluşturulacaktır. Bu çalışmaların, nesli tükenmek üzere olan kuş türlerinin belirlenmesi ve koruma altına alınması için fayda sağlayacağına inanılmaktadır.

IV. KAYNAKLAR

- [1] A. Thakur, V. Abrol, P. Sharma, and P. Rajan, “Local Compressed Convex Spectral Embedding for Bird Species Identification,” *The Journal of the Acoustical Society of America*, c. 143(6), ss. 3819-3828, 2018.

- [2] Xeno-Canto Veri Seti, May. 24, 2019. [Çevrimiçi]. Erişim Adresi: <https://www.xeno-canto.org/>.
- [3] W. Chu, and D.T. Blumstein, "Noise Robust Bird Song Detection Using Syllable Pattern-Based Hidden Markov Models," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011, ss. 345-348.
- [4] J. A. Kogan and D. Margoliash, "Automated Recognition of Bird Song Elements from Continuous Recordings Using Dynamic Time Warping and Hidden Markov Models: A Comparative Study," *The Journal of the Acoustical Society of America*, 1998, c. 103, s. 4, ss. 2185–2196.
- [5] A. Marini, A. J. Turatti, A. S. Britto, and A. L. Koerich, "Visual Andacoustic Identification of Bird Species," *2015 IEEE International Conference on Acoustics, Speech and Signal Processing* 2015, ss. 2309–2313.
- [6] A. L. McIlraith and H. C. Card, "Birdsong Recognition with DSP and Neural Networks," *IEEE WESCANEX 95. Communications, Power, and Computing. Conference Proceedings*, 1995, c. 2, ss. 409–414.
- [7] A. L. McIlraith and H. C. Card, "Birdsong Recognition Using Backpropagation and Multivariate Statistics," *IEEE Transactions on Signal Processing*, c. 199745(11), ss. 2740–2748.
- [8] D. Chakraborty, P. Mukker, P. Rajan, and A. Dileep, "Bird Call Identification Using Dynamic Kernel Based Support Vector Machines and Deep Neural Networks," in *Proceedings of Int. Conf. Mach. Learn. App.* 2016, ss. 280–285.
- [9] E. M. Albornoz, L. D. Vignolo, J. A. Sarquis, and E. Leon, "Automatic Classification of Furnariidae Species from the Paranaense Littoral Region Using Speech-Related Features and Machine Learning," *Ecological Informatics*, 2017, c. 38, ss. 39–49.
- [10] N. Priyadarshani, S. Marsland, J. Juodakis, I. Castro, and V. Listanti, "Wavelet Filters for Automated Recognition of Birdsong in Long-Time Field Recordings," *Methods in Ecology and Evolution*, 2020, c. 11(3), ss. 403-417.
- [11] D. E. Kroodsma, E. H. Miller, and H. Ouellet, "Acoustic Communication in Birds: Song Learning and Its Consequences," *Academic*, New York, 1982, c. 2.
- [12] Á. Incze, H. B. Jancsó, Z. Szilágyi, A. Farkas, and C. Sulyok, "Bird Sound Recognition Using a Convolutional Neural Network," *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics*, 2018, ss. 295-300.
- [13] J. P. Campbell, "Speaker Recognition: A Tutorial," *Proceedings of the IEEE*, 1997, c. 85(9), ss. 1437-1462.
- [14] J. R. Deller, J. H. L. Hansen, & J. G. Proakis, "Discrete-Time Processing of Speech Signals," IEEE Press, Piscataway, N.J, 2000.
- [15] K. Hechenbichler and K. Schliep, "Weighted K-Nearest-Neighbor Techniques and Ordinal Classification Technical Report," Ludwig-Maximilians-Universität München, Institut für Statistik, 2004.
- [16] J. Koronacki and J. Cwik, "Statistical Learning Systems (in Polish)," Wydawnictwa Naukowo-Techniczne, Warsaw, Poland, 2005.

- [17] Hyeran Byun and Seong-Whan Lee, “Applications of Support Vector Machines for Pattern Recognition: A Survey,” *In Proceedings of the First International Workshop on Pattern Recognition with Support Vector Machines, SVM '02*, London, UK, 2002, ss. 213–236.
- [18] C.J.C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” *Data Min. Knowl. Discov.*, 1998, c. 2(2), ss. 121–167.
- [19] Ö. Aydemir, “Ardışıl İleri Yönlü Öznitelik Seçim Algoritmasında Etkin Özniteliklerin Belirlenmesi,” *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, c. 8(3), ss. 495-501, 2017.